

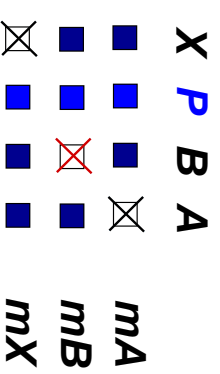
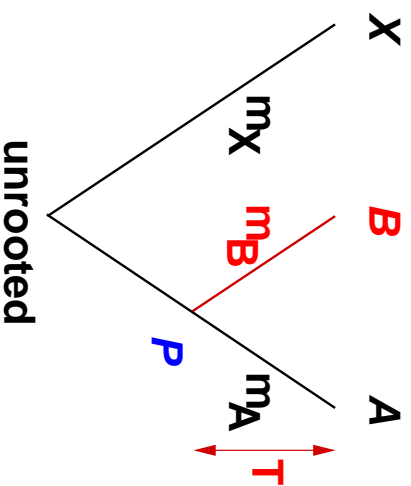
**Relative Rate Tests**  
**Accelerated Evolution Rates in Coding and**  
**Non-Coding Sequences**

Sonja J. Prohaska

*Bioinformatik, Inst.f.Informatik, Univ.Leipzig, Germany*

## Existing Tests for Equality of Evolutionary Rates

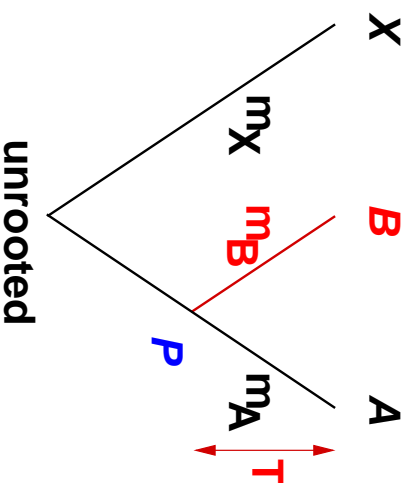
### Tajima test



- $m_B$  = characters exclusively shared between X and A
- $m_A$  = characters exclusively shared between X and B

## Existing Tests for Equality of Evolutionary Rates

### Tajima test



$$E(m_A) = E(m_B)$$

$$\chi^2 = \frac{(m_A - m_B)^2}{m_A + m_B}$$

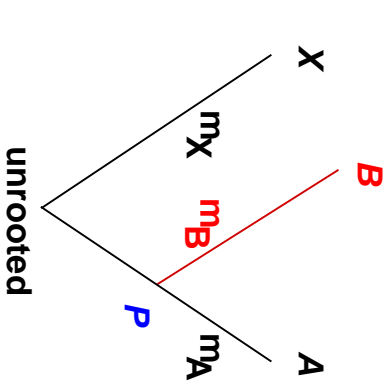
significant difference at 5% level if

$$\chi^2 > 3.841$$

$m_B$  = characters exclusively shared between X and A

$m_A$  = characters exclusively shared between X and B

## Existing Tests for Equality of Evolutionary Rates

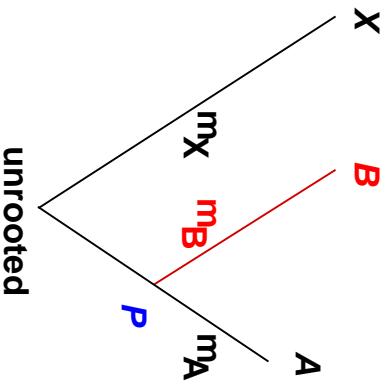


*rate of B > rate of A*

accelerated rate of B?

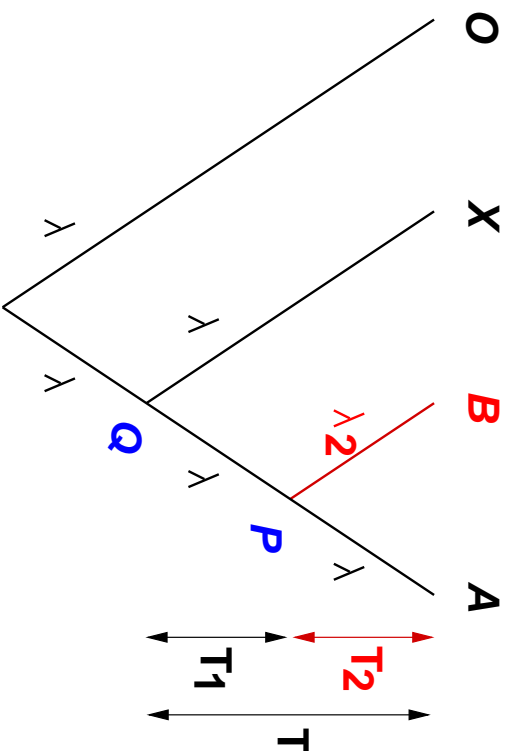
Or

decelerated rate of A?



## New Test for Relative Rates

### Wagner's Test

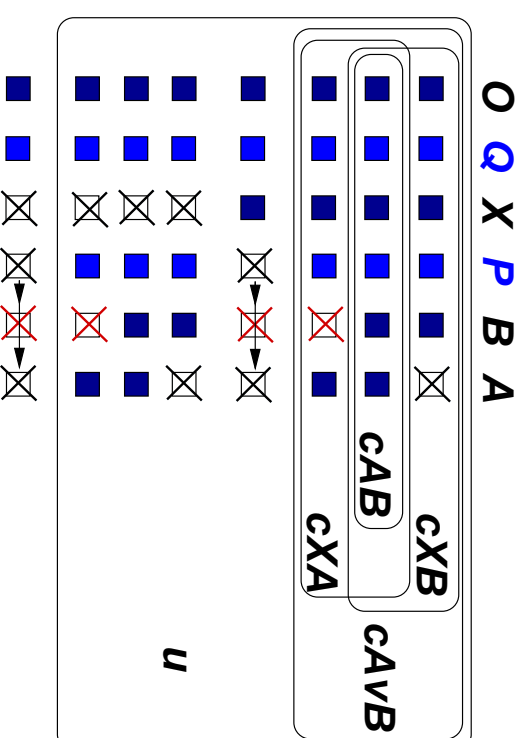
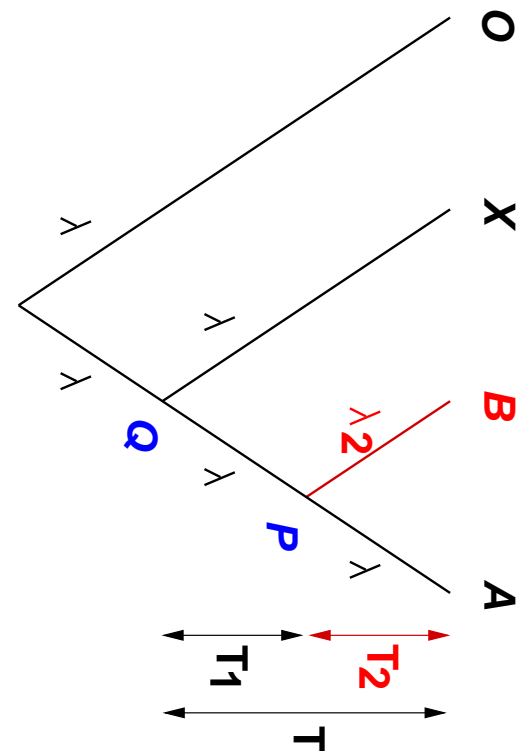


- uses additional outgroup  $O$  to estimate the conserved nucleotides at  $Q$
- assumes exponential decay law for the retention of conserved nucleotides
- tests the hypothesis  $\lambda = \lambda_2$

especially designed to compare evolution rates in non-coding sequences  
accelerated evolution rates can be due to positive selection

## New Test for Relative Rates

### Wagner's Test

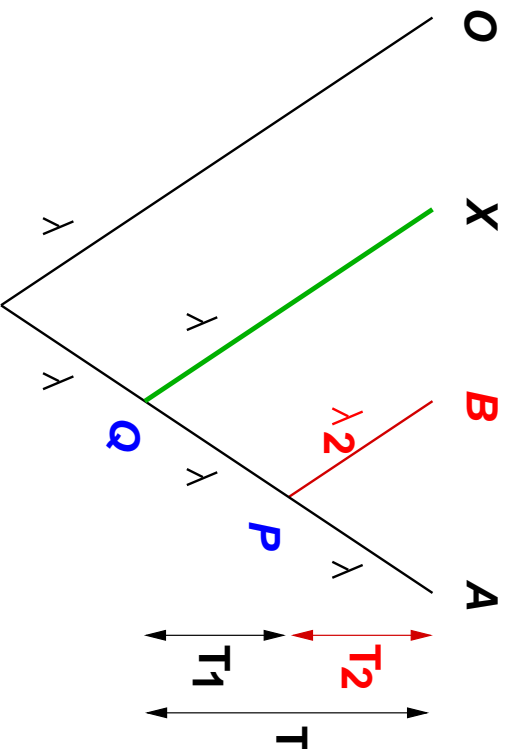


$m'_A = c_{XA} - c_{AB} =$  characters exclusively shared between X and A

$m'_B = c_{XB} - c_{AB} =$  characters exclusively shared between X and B

## New Test for Relative Rates

### Wagner's Test



$$r(QX) = e^{-\lambda T}$$

$$r(QP) = e^{-\lambda T_1}$$

$$r(PA) = e^{-\lambda T_2}$$

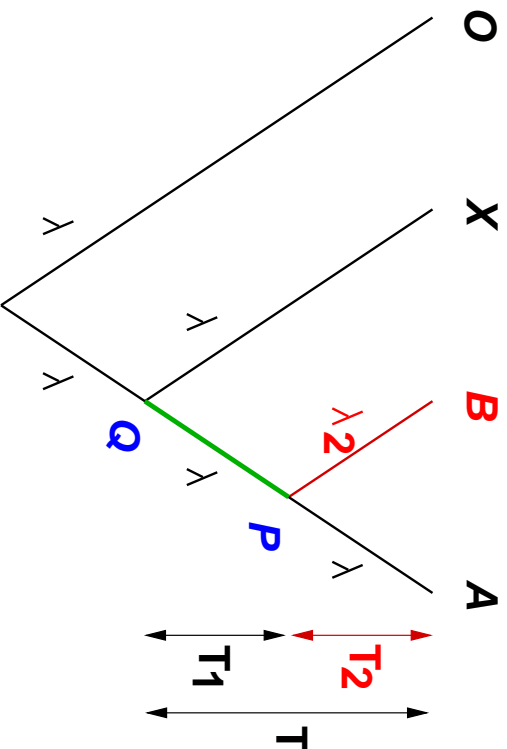
$$r(PB) = e^{-\lambda_2 T_2}$$

$$c_{AB/q} = e^{-\lambda T} e^{-\lambda T_1} e^{-\lambda T_2} e^{-\lambda_2 T_2}$$

calculate the retention ( $c_{AB/q}$ ) of CNS in X, B and A, given the set of CNS in Q

## New Test for Relative Rates

### Wagner's Test



$$r(QX) = e^{-\lambda T}$$

$$r(QP) = e^{-\lambda T_1}$$

$$r(PA) = e^{-\lambda T_2}$$

$$r(PB) = e^{-\lambda 2T_2}$$

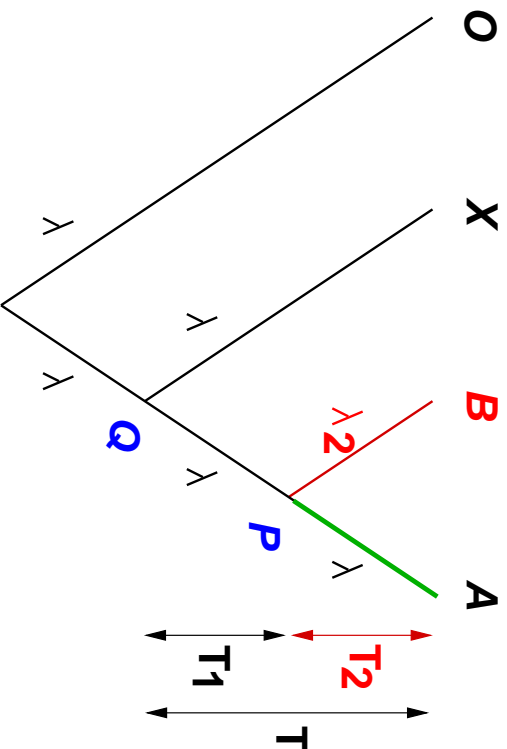
$$c_{AB/q} = e^{-\lambda T} e^{-\lambda T_1} e^{-\lambda T_2} e^{-\lambda 2T_2}$$

calculate the retention ( $c_{AB/q}$ ) of CNS in X, B and A, given the set of CNS in Q



## New Test for Relative Rates

### Wagner's Test



$$r(QX) = e^{-\lambda T}$$

$$r(QP) = e^{-\lambda T_1}$$

$$r(PA) = e^{-\lambda T_2}$$

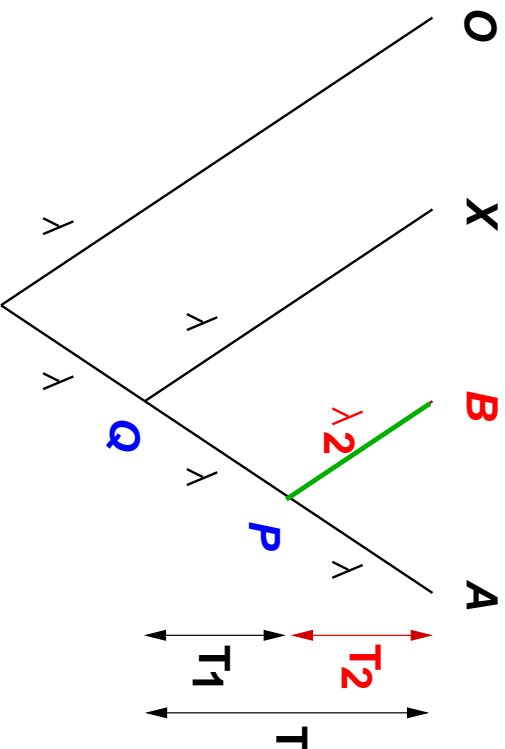
$$r(PB) = e^{-\lambda_2 T_2}$$

$$c_{AB/q} = e^{-\lambda T} e^{-\lambda T_1} e^{-\lambda T_2} e^{-\lambda_2 T_2}$$

calculate the retention ( $c_{AB/q}$ ) of CNS in X, B and A, given the set of CNS in Q

## New Test for Relative Rates

### Wagner's Test



$$r(QX) = e^{-\lambda T}$$

$$r(QP) = e^{-\lambda T_1}$$

$$r(PA) = e^{-\lambda T_2}$$

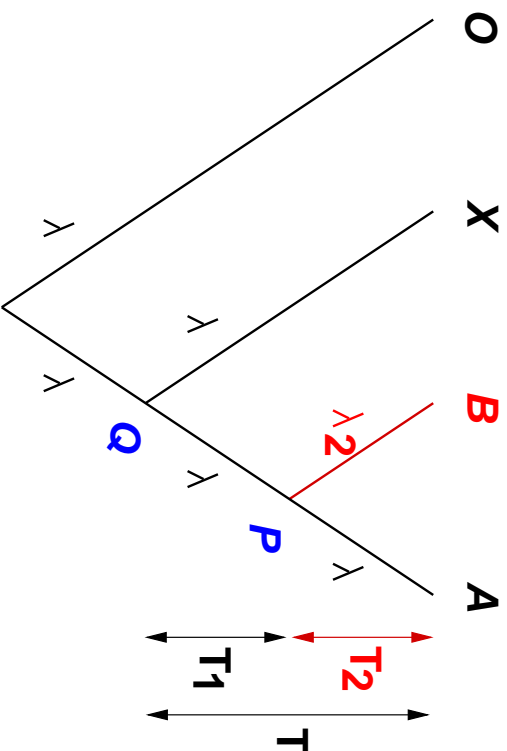
$$r(PB) = e^{-\lambda_2 T_2}$$

$$c_{AB/q} = e^{-\lambda T} e^{-\lambda T_1} e^{-\lambda T_2} e^{-\lambda_2 T_2}$$

calculate the retention ( $c_{AB/q}$ ) of CNS in X, B and A, given the set of CNS in Q

## New Test for Relative Rates

### Wagner's Test



$$c_{AB}/q = e^{-\lambda T} e^{-\lambda T_1} e^{-\lambda T_2} e^{-\lambda_2 T_2}$$

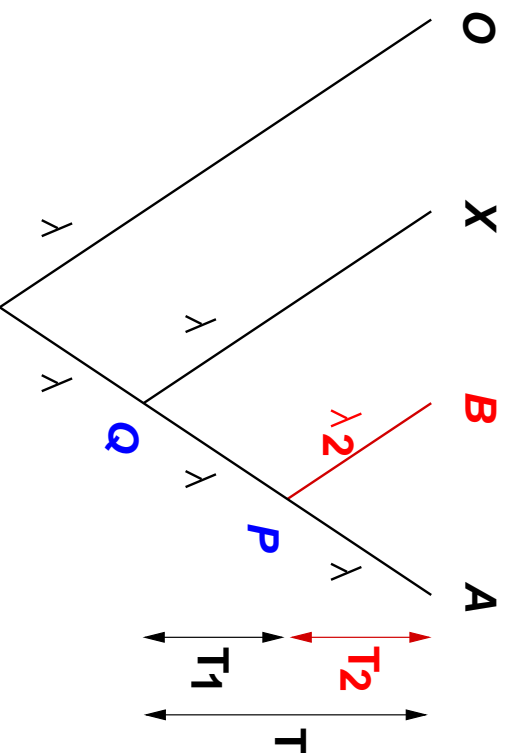
$$c_{XA}/q = e^{-\lambda T} e^{-\lambda T}$$

$$c_{XB}/q = e^{-\lambda T} e^{-\lambda T_1} e^{-\lambda_2 T_2}$$

$$u/q = 1 - (1 - e^{-\lambda T}) \left[ (1 - e^{-\lambda T_1}) + e^{-\lambda T_1} (1 - e^{-\lambda T_2}) (1 - e^{-\lambda_2 T_2}) \right]$$

## New Test for Relative Rates

### Wagner's Test

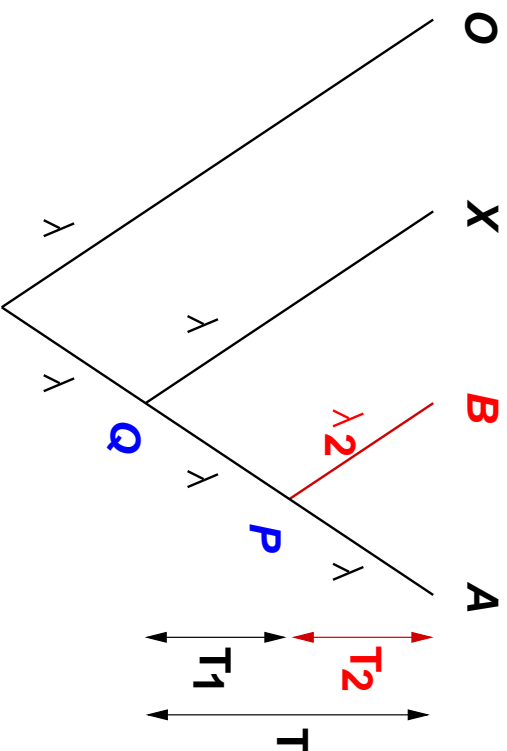


the amount of CNS  
retained in  $Q$  is

$$q = c_{XA} \left( \frac{u + c_{AVB}}{c_{XA} + c_{AVB}} \right)^2$$

## New Test for Relative Rates

### Wagner's Test



variance of an exponential process

$$s^2 = be^{-at}(1 - e^{-at})$$

$$\sigma^2 = 2qe^{-\lambda T_1} e^{-\lambda T_2}(1 - e^{-\lambda T_2})$$

$$2c_{XA} \left( \frac{u + c_{AVB}}{c_{XA} + c_{AVB}} \right) \left( 1 - \frac{c_{AB}}{c_{XB}} \right)$$

## Test Statistics for conserved non-coding nucleotide (CNCN) sequence elements

### Wagner's Test - test statistic

$$z' = \frac{m'_A - m'_B}{\sqrt{\bar{\ell}\sigma^2}} \quad \text{normal distributed with variance 1}$$

- $m'_A$  ...number of CNS exclusively lost along  $PA$
- $m'_B$  ...number of CNS exclusively lost along  $PB$
- $\sigma^2$  ...variance
- $\bar{\ell}$  ...average length of a contiguous CN sequence element

To account for **stochastic dependency** among CNS we scale the sampling variance with  $\bar{\ell}$ .